Numerical Experiments on The Capacity of Quantum Channel with Entangled Input States

Susumu Osawa
High Energy Accelerator Research Organization (KEK),
Tsukuba, Ibaraki 305-0801, Japan.
E-mail: osawa@post.kek.jp
and
Hiroshi Nagaoka
Graduate School of Information Systems,
University of Electro-Communications,
Chofu, Tokyo 182–8585, Japan.
E-mail: nagaoka@is.uec.ac.jp

SUMMARY The capacity of quantum channel with product input states was formulated by the quantum coding theorem. However, whether entangled input states can enhance the quantum channel is still open. It turns out that this problem is reduced to a special case of the more general problem whether the capacity of product quantum channel exhibits additivity. In the present study, we apply one of the quantum Arimoto-Blahut type algorithms to the latter problem. The results suggest that the additivity of product quantum channel capacity always holds and that entangled input states cannot enhance the quantum channel capacity [1].
key words: quantum entanglement, quantum channel capacity, quantum coding theorem, quantum information theory

1 Introduction

The coding theorem for quantum channels[2] was proved in recent publications [11, 21] combined with pioneering works such as [9, 10]. This theorem gives the formula of the capacity of quantum channel with product (not entangled) input states. On the other hand, the use of entangled states in quantum communications provides us with another interesting aspect, which was already pointed out in [3]. Even though lots of theoretical attempts have been made so far in this direction (see section 2.4 for recent works on this subject), we are still far from deep understanding of the role of entanglement in quantum communications. In particular, whether the use of entangled input states can increase the capacity of quantum channel is a big open problem, which, as is seen from Theorem 1 below, can be reduced to a special case of another open problem whether the capacity of product quantum channel exhibits additivity. In the present study, we examine the latter

[1] The content of this paper was partly presented at the second QIT [19] and the 22nd SITA [20].
[2] In this paper we treat only quantum memoryless channels and simply call them quantum channels.

problem numerically by means of a quantum version of Arimoto-Blahut algorithm [18], and observe that the additivity seems to always hold.

2 Capacity of quantum channel and entangled states

2.1 Quantum channel with product input states

In this section, we give a brief review of the standard notion of quantum channel with product input states and its capacity.

Let $\mathcal{H}$ be a Hilbert space which corresponds to a quantum system. A quantum state is represented by a density operator on $\mathcal{H}$, i.e. non-negative operator with unit trace. We denote by $\mathcal{S}(\mathcal{H})$ the totality of density operators on $\mathcal{H}$. Letting $\mathcal{H}_1$ and $\mathcal{H}_2$ be input and output systems, a quantum channel is described by a completely positive [23] trace preserving linear map

$$\Gamma \ : \ \mathcal{T}(\mathcal{H}_1) \to \mathcal{T}(\mathcal{H}_2)$$

where $\mathcal{T}(\mathcal{H}_1)$ and $\mathcal{T}(\mathcal{H}_2)$ are the totalities of the trace class operators on $\mathcal{H}_1$ and $\mathcal{H}_2$. Note that the complete positivity and the trace preservation jointly characterize the physical realizability of quantum channels [17].

A quantum communication system in which a quantum channel Γ is used n times is described as follows. A message set $\mathcal{M}_n := \{1, 2, \cdots, M_n\}$ denotes the totality of the messages which are to be transmitted. Each message $k \in \mathcal{M}_n$ is encoded to a codeword which is a product state in the form $\rho^{(n)}(k) := \rho_1(k) \otimes \cdots \otimes \rho_n(k)$ on $\mathcal{H}_1^{\otimes n}$, where $\mathcal{H}_1^{\otimes n}$ denotes the tensor product Hilbert space $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_1$. The sender transmits the codeword by multiple use of a quantum channel Γ. Then the received state is also a product state $\Gamma^{\otimes n}(\rho^{(n)}(k)) = \Gamma(\rho_1(k)) \otimes \cdots \otimes \Gamma(\rho_n(k))$ on $\mathcal{H}_2^{\otimes n}$. Here $\Gamma^{\otimes n}$ denotes the n-fold tensor product channel $\Gamma \otimes \cdots \otimes \Gamma$ acting on $\mathcal{T}(\mathcal{H}_1^{\otimes n})$. The receiver estimates which codeword has been actually transmitted by performing an $\mathcal{M}_n$-valued measurement. Mathematically, this measurement is represented by a positive operator valued measure (POVM) $X^{(n)} = \{X_1^{(n)}, \cdots, X_{M_n}^{(n)}\}$ on $\mathcal{H}_2^{\otimes n}$, i.e. $X_k^{(n)} \geq 0 \, (k = 1, \cdots, M_n)$ and $\sum_{k=1}^{M_n} X_k^{(n)} = I$, where I denotes the identity operator on $\mathcal{H}_2^{\otimes n}$.

Given a *coding system* Φ_n consisting of codewords $\{\rho^{(n)}(k)\}_{k=1}^{M_n}$ and a measurement $X^{(n)}$, the error probability averaged over all codewords is given by

$$P_{er}(\Phi_n, \Gamma) = 1 - \frac{1}{M_n} \sum_{k=1}^{M_n} \mathrm{Tr}\, [\Gamma^{\otimes n}(\rho^{(n)}(k)) X_k^{(n)}], \tag{1}$$

and the quantity $R_n(\Phi_n) := \log M_n / n$ is called the *rate* of the coding system Φ_n. Now the capacity of the quantum channel Γ with product input states is defined as

$$C(\Gamma) := \sup_{\{\Phi_n\}} \{ \lim_{n \to \infty} R_n(\Phi_n) \ ; \ \lim_{n \to \infty} P_{er}(\Phi_n, \Gamma) = 0 \}. \tag{2}$$

Next let us introduce the quantum mutual information. Let

$$\begin{aligned} \Pi_n \ &:= \ \{(\lambda_1, \cdots, \lambda_n \, ; \, \sigma_1, \cdots, \sigma_n) \ ; \\ &\qquad 0 \leq \lambda_i \in \mathbf{R}, \ \sum_{i=1}^{n} \lambda_i = 1, \ \sigma_i \in \mathcal{S}(\mathcal{H}_1)\}, \\ \Pi \ &:= \ \bigcup_{n}^{\infty} \Pi_n. \end{aligned}$$

An element $\pi = (\lambda_1, \cdots, \lambda_n\,;\, \sigma_1, \cdots, \sigma_n) \in \Pi$ is considered as a discrete probability distribution on $\mathcal{S}(\mathcal{H}_1)$ assigning probability λ_i to the state σ_i for each i. The quantum mutual information for π and Γ is then defined as

$$\mathrm{I}(\pi;\Gamma) := \sum_j \lambda_j \mathrm{D}(\Gamma(\sigma_j)\|\Gamma(\rho)), \tag{3}$$

where $\rho := \sum_j \lambda_j \sigma_j$ is a convex combination of the states in π and $\mathrm{D}(\rho\|\sigma) := \mathrm{Tr}\,[\rho(\log\rho - \log\sigma)]$ is the quantum relative entropy.

Now, the quantum channel coding theorem [9, 10, 11, 21] states that

$$C(\Gamma) = \sup_{\pi\in\Pi} \mathrm{I}(\pi;\Gamma). \tag{4}$$

In addition, supremization is reduced to maximization on certain finite-dimensional compact set [7]. That is,

$$C(\Gamma) = \max_{\pi\in\Pi_n} \mathrm{I}(\pi;\Gamma) = \max_{\pi\in\Pi_n^e} \mathrm{I}(\pi;\Gamma)$$

where $n = \dim \Gamma(\mathcal{S}(\mathcal{H}_1)) + 1$, $\Pi_n^e := \{(\lambda_i;\sigma_i) \in \Pi_n\ ;\ \sigma_i \in \partial_e\mathcal{S}(\mathcal{H}_1),\, i = 1, \cdots, n\}$. Here $\partial_e\mathcal{S}(\mathcal{H}_1)$ is the totality of extreme points (pure states) of $\mathcal{S}(\mathcal{H}_1)$.

2.2 Quantum channel with entangled input states

Some states on a tensor product Hilbert space cannot be represented as product states or their convex combinations. These states are called *entangled states*. In the formulation given in the previous section we treated only product states as inputs to (and, consequently, outputs from) a quantum channel. Now let us consider communication systems in which we are allowed to use entangled input states.

The capacity of the quantum channel Γ with entangled input states $\tilde{C}(\Gamma)$ is defined in the same way as (2) except that arbitrary states on $\mathcal{H}_1^{\otimes n}$, not necessarily product states, can be codewords. It is obvious by definition that

$$\tilde{C}(\Gamma) \geq C(\Gamma). \tag{5}$$

However, neither example of channel exhibiting the strict inequality nor proof that the equality

$$\tilde{C}(\Gamma) = C(\Gamma) \tag{6}$$

always holds have been reported yet[3]. This problem can be reduced to the additivity problem for the capacity of product channels as described in the next section.

2.3 Capacity of product quantum channel

Let $\Gamma^{(i)} : \mathcal{T}(\mathcal{H}_1^{(i)}) \to \mathcal{T}(\mathcal{H}_2^{(i)})$ for $i = 1, 2$ be quantum channels, and let $\Gamma^{(1)} \otimes \Gamma^{(2)} : \mathcal{T}(\mathcal{H}_1^{(1)} \otimes \mathcal{H}_1^{(2)}) \to \mathcal{T}(\mathcal{H}_2^{(1)} \otimes \mathcal{H}_2^{(2)})$ be their product channel. The capacity $C(\Gamma^{(1)} \otimes \Gamma^{(2)})$ is defined as (2) by replacing Γ with $\Gamma^{(1)} \otimes \Gamma^{(2)}$ in which each input state (code word) is written in the form $\rho^{(n)}(k) = \rho_1(k) \otimes \cdots \otimes \rho_n(k)$, where $\rho_i(k)$ $(i = 1, \cdots, n)$ are arbitrary states on $\mathcal{H}_1^{(1)} \otimes \mathcal{H}_1^{(2)}$. Then it is easy to see that the superadditivity

$$C(\Gamma^{(1)} \otimes \Gamma^{(2)}) \geq C(\Gamma^{(1)}) + C(\Gamma^{(2)}) \tag{7}$$

[3] An earlier example of statement of this problem is found in the concluding remarks of [7].

holds. However, as in the case of (5), we have no example of the strict inequality nor proof that the additivity

$$C(\Gamma^{(1)} \otimes \Gamma^{(2)}) = C(\Gamma^{(1)}) + C(\Gamma^{(2)}) \tag{8}$$

always holds. Actually, this problem includes the previous one as is seen from the following theorem, whose proof is given in Appendix A.

Theorem 1:

$$\tilde{C}(\Gamma) = \lim_{N\to\infty} \frac{C(\Gamma^{\otimes N})}{N} = \sup_N \frac{C(\Gamma^{\otimes N})}{N} \tag{9}$$

holds. Here $C(\Gamma^{\otimes N})$ is defined as (2) by replacing Γ with $\Gamma^{\otimes N}$ in which each input state is written in the form $\rho^{(n)}(k) = \rho_1(k) \otimes \cdots \otimes \rho_n(k)$, where $\rho_i(k)$ $(i = 1, \cdots, n)$ are arbitrary states on $\mathcal{H}_1^{\otimes N}$.

If the additivity (8) always holds, we have $C(\Gamma^{\otimes N}) = NC(\Gamma)$, which, combined with Theorem 1, leads to the equality (6). In other words, the additivity implies that entanglement of input states cannot increase the capacity of quantum channel.

2.4 The aim of the present paper and related works

In the last few years the additivity (8) has gradually been receiving recognition as a difficult but important problem in the quantum information theory, and has been proved for several special cases. At the moment, proofs are known for the cases when $\Gamma^{(1)}$ is arbitrary and $\Gamma^{(2)}$ is the identity [22], when $\Gamma^{(1)}$ is arbitrary and $\Gamma^{(2)}$ is either a Holevo's classical-quantum or quantum-classical channel [15] and when $\Gamma^{(1)}$ is arbitrary and $\Gamma^{(2)}$ is a certain class of unital binary channels [16]. See also [5, 12], whose results are now included in some of the above-mentioned ones. On the other hand, no example of channel violating the additivity have been found so far, and naturally the conjecture that the additivity always holds is arising [1, 12, 14, 15, 19, 20, 22]. The aim of the present paper is to show that an efficient algorithm for computing quantum channel capacity, which was recently introduced by one of the authors, is applicable to verification of the conjecture and to report that all the randomly chosen channels have exhibited the additivity [4].

3 Numerical experiments on the additivity

3.1 Quantum version of Arimoto-Blahut algorithm

The Arimoto-Blahut algorithm is known for computing the capacity of classical channel [2, 4]. Recently, one of the authors proposed some algorithms of this type for computing the capacity of quantum channel [18, 19]. We use one of these, which is called the *boundary algorithm* since its recursion works on the extreme boundary set $\partial_e \mathcal{S}(\mathcal{H}_1)$. The outline of the theoretical basis is as follows.

[4] In several references such as [1, 13, 14] it is shortly mentioned, without any detail, that some numerical works relating the conjecture have been made.

Let us introduce a two-variable extension of $\mathrm{I}(\pi;\Gamma)$:

$$J(\pi,\pi') := -\mathrm{D}(\lambda\|\lambda') + \sum_{i=1}^{n} \lambda_i \mathrm{Tr}\left[\Gamma(\sigma_i)\Phi(\sigma_i',\rho')\right], \tag{10}$$

where

$$\pi = (\lambda_i;\sigma_i),\ \pi' = (\lambda_i';\sigma_i') \in \Pi_n,$$
$$\mathrm{D}(\lambda\|\lambda') := \sum_{i=1}^{n} \lambda_i \log\frac{\lambda_i}{\lambda_i'},\ \rho' := \sum_{i=1}^{n} \lambda_i'\sigma_i',$$
$$\Phi(\sigma_i',\rho') := \log(\Gamma(\sigma_i')) - \log(\Gamma(\rho')).$$

Then it holds that

$$\mathrm{I}(\pi;\Gamma) = J(\pi,\pi) = \max_{\pi'} J(\pi,\pi'). \tag{11}$$

In addition, we can compute $\hat{\pi} = (\hat{\lambda_i},\, \hat{\sigma_i}) := \mathrm{argmax}_{\pi}\, J(\pi,\pi')$ by the following equations.

$$\hat{\sigma_i} = \operatorname*{argmax}_{\sigma\in\mathcal{S}(\mathcal{H}_1)} \mathrm{Tr}\left[\Gamma(\sigma)\Phi(\sigma_i',\rho')\right],$$

$$\hat{\lambda_i} = \lambda_i' \exp(\mathrm{Tr}\left[\Gamma(\hat{\sigma_i})\Phi(\sigma_i',\rho')\right])/\hat{Z},$$

where $\hat{Z}$ is the normalizing constant:

$$\hat{Z} := \sum_{i=1}^{n} \lambda_i' \exp(\mathrm{Tr}\left[\Gamma(\hat{\sigma_i})\Phi(\sigma_i',\rho')\right]).$$

Note that, since $\mathrm{Tr}\left[\Gamma(\sigma)\Phi(\sigma_i',\rho')\right]$ is linear in σ, we can always choose $\hat{\sigma_i}$ to be an extreme point of $\mathcal{S}(\mathcal{H}_1)$, i.e. a pure state $|\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle$ is a normalized eigenvector of $\Gamma^*(\Phi(\sigma_i',\rho'))$ corresponding to the maximum eigenvalue. Here $\Gamma^* : \mathcal{B}(\mathcal{H}_2) \to \mathcal{B}(\mathcal{H}_1)$ denotes the dual map of Γ defined by $\mathrm{Tr}\left[\Gamma(X)Y\right] = \mathrm{Tr}\left[X\Gamma^*(Y)\right]$ for $\forall X \in \mathcal{T}(\mathcal{H}_1)$ and $\forall Y \in \mathcal{B}(\mathcal{H}_2)$, where $\mathcal{B}(\mathcal{H}_i)$ $(i = 1, 2)$ are the totalities of bounded operators on $\mathcal{H}_i$.

Given a number $n \leq \dim\Gamma(\mathcal{S}(\mathcal{H}_1)) + 1$ and an arbitrary initial element $\pi^{(1)} \in \Pi_n$, let the sequence $\{\pi^{(k)}\}_{k=1}^{\infty}$ be defined by

$$\pi^{(k+1)} := \operatorname*{argmax}_{\pi} J(\pi,\pi^{(k)}). \tag{12}$$

Note that the sequence $\{\mathrm{I}(\pi^{(k)};\Gamma)\}_{k=1}^{\infty}$ is monotonous, since

$$\mathrm{I}(\pi^{(k)};\Gamma) \leq J(\pi^{(k+1)},\pi^{(k)}) \leq \mathrm{I}(\pi^{(k+1)};\Gamma) \tag{13}$$

holds. Therefore we can efficiently compute the limit value $\lim_{k\to\infty}\mathrm{I}(\pi^{(k)};\Gamma)$. Unfortunately, it is not necessarily the quantum channel capacity since the quantum version of Arimoto-Blahut algorithm does not assure the global maximum. Thus we make several convergent sequences and adopt the maximum limit value as an estimate of the capacity. We judge that a sequence reaches the limit value when ten successive numerical values are the same to six places of decimals.

Table 1: Examples of quantum binary channels (A, b)

	A	b
Γ_1	$\begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.2 \end{pmatrix}$	$\begin{pmatrix} 0.2 \\ 0 \\ 0 \end{pmatrix}$
Γ_2	$\begin{pmatrix} 0.05 & -0.2 & 0.4 \\ -0.2 & -0.05 & -0.2 \\ 0.2 & 0 & -0.5 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0.1 \end{pmatrix}$
Γ_3	$\begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix} \cdot \begin{pmatrix} -0.45 & 0 & 0 \\ 0 & 0.6 & 0 \\ 0 & 0 & -0.6 \end{pmatrix} \cdot \begin{pmatrix} 0.8 & 0.6 & 0 \\ 0.6 & -0.8 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.2 \\ -0.2 \\ 0.2 \end{pmatrix}$
Γ_4	$\begin{pmatrix} 0.1 & -0.3 & 0 \\ -0.3 & -0.1 & -0.2 \\ 0 & 0 & -0.05 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.2 \\ 0.55 \end{pmatrix}$

Table 2: Examples of quantum channels with $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^3$ having generators of the form $\{V_1, V_2, \sqrt{I - V_1^* V_1 - V_2^* V_2}\}$

	V_1	V_2
Γ_5	$\begin{pmatrix} 0.2 & 0.3 & 0.4 \\ 0 & 0.5i & 0 \\ 0.1i & 0.4i & 0.5i \end{pmatrix}$	$\begin{pmatrix} 0.1 - 0.3i & 0 & 0 \\ 0 & -0.3i & 0.1 - 0.2i \\ 0.3 - 0.3i & 0.2 + 0.1i & 0 \end{pmatrix}$
Γ_6	$\begin{pmatrix} 0.19 & 0.7 & -0.1 + 0.3i \\ 0.4i & 0.06 & -0.1 + 0.05i \\ 0.2 & 0.39 & 0.4 - 0.4i \end{pmatrix}$	$\begin{pmatrix} 0.3 & -0.1 & 0.1 \\ 0.2 & 0.3 & 0.02i \\ 0.1 & 0.2 & 0.1i \end{pmatrix}$

3.2 Setting of the experiments

We apply the quantum Arimoto-Blahut algorithm to various quantum channels $\Gamma^{(1)}$ and $\Gamma^{(2)}$ as well as their product channels $\Gamma^{(1)} \otimes \Gamma^{(2)}$ to examine whether the additivity (8) always holds. Here we restrict ourselves to the case when $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^2$ or $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^3$ to reduce the computational complexity. Tables 1 and 2 show representative examples of channels used as $\Gamma^{(1)}$ and $\Gamma^{(2)}$ in the experiments. Here the first four channels $\Gamma_1, \cdots, \Gamma_4$ are quantum binary channels in the sense that $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^2$, and are expressed in terms of the coefficients (A, b) of the corresponding affine transformations on $\mathbf{R}^3$ (see Appendix C), while the other channels Γ_5 and Γ_6 are of $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^3$ and are expressed by the generators $\{V_1, \cdots, V_m\}$ of their operator-sum representations (see Appendix B), restricting ourselves to the case when $m = 3$ and $V_3 = \sqrt{I - V_1^* V_1 - V_2^* V_2}$. Note that these channels do not belong to the special classes mentioned in section 2.4 for which the additivity has been proved. These examples are chosen basically in random manners so that they are as generic as possible, not being intended to have any special properties or to represent any concrete physical processes, except that some extra points are taken into consideration in view of computational efficiency and generality, as explained below.

In the case of quantum channels with $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^2$, the necessary and sufficient condition for the coefficients (A, b) to represent a pseudoclassical channel is known [7]. Here a quantum channel is said to be *pseudoclassical* when its capacity is unchanged even

Table 3: The capacity of the quantum channels shown in Tables 1 and 2 and their product channels

$\Gamma^{(1)}$	$\Gamma^{(2)}$	$C(\Gamma^{(1)})$	$C(\Gamma^{(2)})$	$C(\Gamma^{(1)})+C(\Gamma^{(2)})$	$C(\Gamma^{(1)}\otimes\Gamma^{(2)})$
Γ_1	Γ_1	0.138166	0.138166	0.276311	0.276311
Γ_2	Γ_2	0.258679	0.258679	0.517358	0.517358
Γ_1	Γ_3	0.138166	0.243068	0.381233	0.381233
Γ_3	Γ_2	0.243068	0.258679	0.501747	0.501746
Γ_2	Γ_4	0.258679	0.0898225	0.348501	0.348501
Γ_5	Γ_5	0.677358	0.677358	1.354716	1.354716
Γ_6	Γ_6	0.829580	0.829580	1.659160	1.659160
Γ_5	Γ_6	0.677358	0.829580	1.506938	1.506938

if the measurements $X^{(n)}$ on $\mathcal{H}_2^{\otimes n}$ in equation (1) are restricted to separable measurements which are constructed from the tensor products of measurements on $\mathcal{H}_2$. Considering the fundamental importance of the pseudoclassicality in classification of quantum channels, we choose Γ_1 and Γ_3 to be pseudoclassical, while Γ_2 and Γ_4 to be non-pseudoclassical, and examine various combinations of these channels.

In the case of quantum channels with $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^3$, on the other hand, we do not care about the pseudoclassicality, since no practical criterion for this property is known. The general operator-sum representation of channel in this case is given by generators $\{V_1, \cdots, V_m\}$ satisfying $\sum_{k=1}^m V_k^* V_k = I$ with $m \leq 9$, while our setting of Γ_5 and Γ_6 is much more restrictive. This restriction simply comes from a demand to reduce computational complexity. Nevertheless, the choice of channels may be considered sufficiently generic in the sense that it does not assume any special structure in view of the additivity.

As we mentioned in section 3.1, the quantum version of Arimoto-Blahut algorithm does not assure the global maximum, and the limit value of $\mathrm{I}(\pi^{(k)}; \Gamma^{(1)})$ or $\mathrm{I}(\pi^{(k)}; \Gamma^{(1)} \otimes \Gamma^{(2)})$ may depend on the initial element $\pi^{(1)} \in \Pi_n$. Therefore, we repeatedly apply the algorithm to a channel with several different initial elements, and adopt the maximum of the convergent values as the estimate of the capacity. However, it empirically appears that the algorithm is not so sensitive to the initial condition. Indeed, we have observed that randomly chosen initial conditions mostly yield the same convergent value as far as the number n of the states in $\pi^{(1)}$ is chosen to be sufficiently large (i.e. $n \approx \dim \Gamma(\mathcal{S}(\mathcal{H}_1))+1$). The following is an example of $\pi^{(1)}$ for which the convergent value has attained the capacity when applied to each of the quantum binary channels $\Gamma_1, \cdots, \Gamma_4$:

$$\pi^{(1)} = (\lambda_1, \cdots, \lambda_4\,;\, \sigma_1, \cdots, \sigma_4)$$

with $\lambda_1 = \cdots = \lambda_4 = 1/4$ and

$$\sigma_1 = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \sigma_2 = \frac{1}{2}\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix},$$

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma_4 = \frac{1}{2\sqrt{3}}\begin{pmatrix} \sqrt{3}-1 & -1-i \\ -1+i & \sqrt{3}+1 \end{pmatrix}.$$

Table 4: The probability distributions which maximize the quantum mutual information of the quantum channels shown in Tables 1 and 2

	$\pi^* = (\lambda_i^*\ ;\ \sigma_i^*)$
Γ_1	$(0.521046, 0.478954\,;\ \begin{pmatrix} 0.500 & 0.500 \\ 0.500 & 0.500 \end{pmatrix}, \begin{pmatrix} 0.500 & -0.500 \\ -0.500 & 0.500 \end{pmatrix})$
Γ_2	$(0.512423, 0.487577\,;\ \begin{pmatrix} 0.009 & 0.055 - 0.079i \\ 0.055 + 0.079i & 0.991 \end{pmatrix}, \begin{pmatrix} 0.991, & -0.049 + 0.082i \\ -0.049 - 0.082i & 0.009 \end{pmatrix})$
Γ_3	$(0.271288, 0.728711\,;\ \begin{pmatrix} 0.00 & 0.00 \\ 0.00 & 1.00 \end{pmatrix}, \begin{pmatrix} 1.00 & 0.00 \\ 0.00 & 0.00 \end{pmatrix})$
Γ_4	$(0.47431, 0.52569\,;\ \begin{pmatrix} 0.772 & 0.398 - 0.133i \\ 0.398 + 0.133i & 0.228 \end{pmatrix}, \begin{pmatrix} 0.218 & -0.392 + 0.131i \\ -0.392 - 0.131 & 0.782 \end{pmatrix})$
Γ_5	$(0.31721, 0.383025, 0.299764\,;\ \begin{pmatrix} 0.690 & -0.365 - 0.260i & 0.111 - 0.034i \\ -0.365 + 0.260i & 0.290 & -0.046 + 0.060i \\ 0.111 + 0.034i & -0.046 - 0.060i & 0.019 \end{pmatrix},$ $\begin{pmatrix} 0.023 & 0.079 + 0.024i & -0.121 - 0.035i \\ 0.079 - 0.024i & 0.294 & -0.448 + 0.004i \\ -0.121 + 0.035i & -0.448 - 0.004i & 0.683 \end{pmatrix}, \begin{pmatrix} 0.157 & 0.273 - 0.004i & 0.233 - 0.055i \\ 0.273 + 0.004i & 0.476 & 0.408 - 0.090i \\ 0.233 + 0.055i & 0.408 + 0.090i & 0.367 \end{pmatrix})$
Γ_6	$(0.327542, 0.285361, 0.387097\,;\ \begin{pmatrix} 0.535 & -0.335 + 0.328i & 0.020 - 0.168i \\ -0.335 - 0.328i & 0.411 & -0.116 + 0.093i \\ 0.020 + 0.168i & -0.116 - 0.093i & 0.054 \end{pmatrix},$ $\begin{pmatrix} 0.392 & 0.375 - 0.250i & 0.152 + 0.113i \\ 0.375 + 0.250i & 0.516 & 0.073 + 0.204i \\ 0.152 - 0.113i & 0.073 - 0.204i & 0.091 \end{pmatrix}, \begin{pmatrix} 0.039 & -0.009 - 0.038i & -0.189 + 0.002i \\ -0.009 + 0.038i & 0.039 & 0.041 - 0.186i \\ -0.189 - 0.002i & 0.041 + 0.186i & 0.922 \end{pmatrix})$

3.3 Results

We have observed that the additivity exactly holds for all the cases we examined, as is seen in Table 3 for the representative examples. Table 4 shows the probability distributions which maximize the quantum mutual information of the quantum channels Γ_i $(i = 1, \cdots, 6)$. In the case of product channels, the probability distribution $\pi^* := \operatorname{argmax}_{\pi} \mathrm{I}(\pi; \Gamma^{(1)} \otimes \Gamma^{(2)})$ has turned out to be the product probability distribution of $\pi_1^* = (\lambda_{i1}^*; \sigma_{i1}^*) := \operatorname{argmax}_{\pi} \mathrm{I}(\pi; \Gamma^{(1)})$ and $\pi_2^* = (\lambda_{j2}^*; \sigma_{j2}^*) := \operatorname{argmax}_{\pi} \mathrm{I}(\pi; \Gamma^{(2)})$, which assigns probability $\lambda_{i1}^* \lambda_{j2}^*$ to the state $\sigma_{i1}^* \otimes \sigma_{j2}^*$.

Fig. 1 illustrates the change in the quantum mutual information $\mathrm{I}(\pi^{(k)}, \Gamma^{(1)} \otimes \Gamma^{(2)})$ for $\Gamma^{(1)} = \Gamma^{(2)} = \Gamma_2$ in the process of recursive computation $\pi^{(k)} \to \pi^{(k+1)}$ starting from some entangled states in $\mathcal{S}(\mathcal{H}_1^{(1)} \otimes \mathcal{H}_1^{(2)})$. In addition, we measure the entanglement [5] of the states in $\pi^{(k)} = (\lambda_i^{(k)}; \sigma_i^{(k)})$ by

$$\mathrm{Ent}(\pi^{(k)}) := \sum_i \lambda_i^{(k)} \mathrm{D}(\sigma_i^{(k)} \| \sigma_{i1}^{(k)} \otimes \sigma_{i2}^{(k)}), \tag{14}$$

where $\sigma_{i1}^{(k)}$ and $\sigma_{i2}^{(k)}$ are the marginal states of $\sigma_i^{(k)}$ defined by partial trace. Fig. 2 shows how the states get disentangled through the recursion.

[5] According to the criterion of [24], the relative entropy $\mathrm{D}(\sigma_i^{(k)} \| \sigma_{i1}^{(k)} \otimes \sigma_{i2}^{(k)})$ is inappropriate as a measure of entanglement in $\sigma_i^{(k)}$ since it takes a positive value even when $\sigma_i^{(k)}$ is a classical mixture (convex combination) of several product states. Nevertheless, it does not mean that its use is inappropriate for our study.

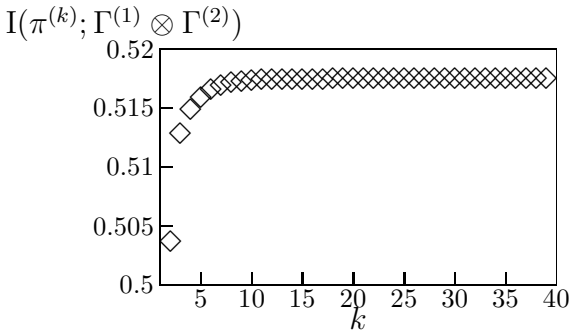

Figure 1: The quantum mutual information $\mathrm{I}(\pi^{(k)}; \Gamma^{(1)} \otimes \Gamma^{(2)})$ in the process of the recursion $\pi^{(k)} \to \pi^{(k+1)}$

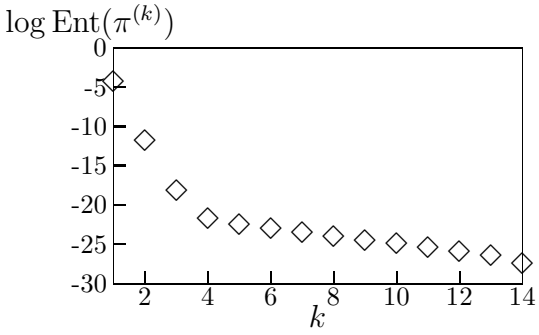

Figure 2: Semi-logarithmic plot of entanglement versus iteration number

4 Conclusions

We have applied the quantum Arimoto-Blahut algorithm to various quantum channels $\Gamma^{(1)}$ and $\Gamma^{(2)}$ with $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^2$ and $\mathcal{H}_1 = \mathcal{H}_2 = \mathbf{C}^3$ as well as their product channels $\Gamma^{(1)} \otimes \Gamma^{(2)}$, and have verified that the additivity (8) holds for all the examples investigated. Note that the additivity (8) has been proved only for some special classes of channels so far as explained in section 2.4 and that the examined channels do not belong to them. Needless to say, this is not a theoretical analysis but a numerical one, applied to only a limited number of channels on low-dimensional Hilbert spaces, using an algorithm which does not ensure the global maximum. Therefore we cannot rely upon the obtained results too much. Nevertheless, it seems very unlikely that all the randomly chosen examples happen to satisfy the additivity by a coincidence or that dimensions 2 and 3 are special in a property like the additivity which is not explicitly related to the dimension. We are thus naturally led to conclude that the results suggest that the additivity always holds.

Acknowledgement

We would like to thank Dr. A.S. Holevo of Steklov Mathematical Institute and Dr. A. Fujiwara of the Osaka University for giving crucial comments on this work.

Appendix

A Proof of Theorem 1

Since $C(\Gamma^{\otimes(N+M)}) \geq C(\Gamma^{\otimes N}) + C(\Gamma^{\otimes M})$ holds, $\lim_{N\to\infty} \frac{C(\Gamma^{\otimes N})}{N}$ exists and is proved to be $\sup_N \frac{C(\Gamma^{\otimes N})}{N}$. Let $\{\Phi_N\}_{N=1}^{\infty}$ be a sequence of coding systems satisfying $\lim_{N\to\infty} P_{er}(\Phi_N, \Gamma) = 0$, where Φ_N consists of codewords $\{\sigma^{(N)}(i)\}_{i=1}^{M_N}$ which are arbitrary states on $\mathcal{H}_1^{\otimes N}$ and a measurement $X^{(N)} = \{X_i^{(N)}\}_{i=1}^{M_N}$. Let Y_N be the classical random variable on the message set $\{1, \cdots, M_N\}$ which takes each value with equal probability $1/M_N$, and let $\hat{Y}_N$ be the classical random variable on the same set which represents the decoded message obtained by performing the measurement $X^{(N)}$ to the output state $\Gamma^{\otimes N}(\sigma^{(N)}(i))$, where the transmitted message i is assumed to be Y_N. Then the Fano inequality (see e.g. [6]) implies

$$1 + P_{er}(\Phi_N, \Gamma) \log M_N \geq \log M_N - \mathrm{I}(Y_N; \hat{Y}_N),$$

where $\mathrm{I}(Y_N; \hat{Y}_N)$ is the classical mutual information between Y_N and $\hat{Y}_N$. This leads to

$$(1 - P_{er}(\Phi_N, \Gamma)) \frac{1}{N} \log M_N \leq \frac{1}{N} + \frac{1}{N} \mathrm{I}(Y_N; \hat{Y}_N). \tag{15}$$

In addition, we have

$$\begin{aligned}
&\mathrm{I}(Y_N; \hat{Y}_N) \\
&\quad = \frac{1}{M_N} \sum_{i=1}^{M_N} \mathrm{D}_{X^{(N)}}(\Gamma^{\otimes N}(\sigma^{(N)}(i)) \| \Gamma^{\otimes N}(\rho^{(N)})) \\
&\quad \leq \frac{1}{M_N} \sum_{i=1}^{M_N} \mathrm{D}(\Gamma^{\otimes N}(\sigma^{(N)}(i)) \| \Gamma^{\otimes N}(\rho^{(N)})) \\
&\quad \leq \max_{\pi} \mathrm{I}(\pi; \Gamma^{\otimes N}) \\
&\quad = C(\Gamma^{\otimes N}),
\end{aligned} \tag{16}$$

where $\mathrm{D}_{X^{(N)}}(\Gamma^{\otimes N}(\sigma^{(N)}(i)) \| \Gamma^{\otimes N}(\rho^{(N)}))$ is the classical relative entropy between the conditional probability $P_{\hat{Y}_N|Y_N}(\cdot|i) := \mathrm{Tr}\,[\Gamma^{\otimes N}(\sigma^{(N)}(i)) X_\cdot^{(N)}]$ and the probability $P_{\hat{Y}}(\cdot) := \mathrm{Tr}\,[\Gamma^{\otimes N}(\rho^{(N)}) X_\cdot^{(N)}]$ with $\rho^{(N)} := \frac{1}{M_N} \sum_i \sigma^{(N)}(i)$, and the first inequality follows from the monotonicity of the relative entropy. Substituting (16) into (15), letting $\mathrm{N} \to \infty$ and taking supremum with respect to $\{\Phi_N\}_{N=1}^{\infty}$, we come to the inequality $\tilde{C}(\Gamma) \leq \lim_{N\to\infty} \frac{C(\Gamma^{\otimes N})}{N}$. Conversely, since $C(\Gamma^{\otimes N})$ is the supremum of the limit values of the rates of asymptotically error-free coding systems whose codewords are restricted to product states of the form $\rho_1 \otimes \cdots \otimes \rho_n \in \mathcal{S}(\mathcal{H}_1^{\otimes Nn})$, where ρ_i $(i = 1, \cdots, n)$ are arbitrary states on $\mathcal{H}_1^{\otimes N}$, it cannot be greater than $N\tilde{C}(\Gamma)$ by the definition of $\tilde{C}(\Gamma)$. Hence we have $\tilde{C}(\Gamma) \geq \lim_{N\to\infty} \frac{C(\Gamma^{\otimes N})}{N}$.

■

B Operator-sum representation

An arbitrary completely positive trace preserving linear map $\Gamma : \mathcal{T}(\mathcal{H}_1) \to \mathcal{T}(\mathcal{H}_2)$ can be written in the form

$$\Gamma(\rho) = \sum_{k=1}^{m} V_k \rho V_k^*$$

where $\mathcal{V} = \{V_k\}_{k=1}^m$ is a collection of bounded operators from $\mathcal{H}_1$ to $\mathcal{H}_2$ satisfying $\sum_{k=1}^m V_k^* V_k = \mathrm{I}$ [17] and m can be taken at most $\dim \mathcal{H}_1 \dim \mathcal{H}_2$ [8]. This is called the *operator-sum representation* or the *Kraus decomposition* of Γ with the *generator* $\mathcal{V}$.

C Quantum binary channel

A quantum channel whose input and output systems are both $\mathbf{C}^2$ is called a *quantum binary channel*. Since every density operator on $\mathbf{C}^2$ is represented in the form

$$\rho_\theta = \frac{1}{2} \begin{pmatrix} 1+\theta_3 & \theta_1 - i\theta_2 \\ \theta_1 + i\theta_2 & 1-\theta_3 \end{pmatrix}$$

with $\theta = (\theta_1, \theta_2, \theta_3)^t$ lying in the unit ball

$$\mathcal{V} = \{\theta \in \mathbf{R}^3 \ ; \ \| \ \theta \ \|^2 = \theta_1^2 + \theta_2^2 + \theta_3^2 \leq 1\},$$

an arbitrary quantum binary channel is represented as $\Gamma(\rho_\theta) = \rho_{A\theta+b}$ by a 3×3 real matrix A and a 3-dimensional real column vector b. We denote such a channel by $\Gamma = (A, b)$. For representing a completely positive map, they should satisfy the following condition [8]

$$\begin{pmatrix} \frac{1}{2}+p & x & r & w \\ \overline{x} & \frac{1}{2}-p & y & -r \\ \overline{r} & \overline{y} & \frac{1}{2}+q & z \\ \overline{w} & -\overline{r} & \overline{z} & \frac{1}{2}-q \end{pmatrix} \geq 0$$

when A and b are represented as

$$A = \begin{pmatrix} y_R + w_R & y_I + w_I & x_R - z_R \\ y_I - w_I & -y_R + w_R & -x_I + z_I \\ 2r_R & 2r_I & p-q \end{pmatrix},$$

$$b = \begin{pmatrix} x_R + z_R \\ -x_I - z_I \\ p+q \end{pmatrix}.$$

(The subscripts R and I denote the real and imaginary parts, i.e. $x = x_R + ix_I$, etc.)

References

[1] G.G. Amosov, A.S. Holevo, and R.F. Werner, "On some additivity problems in quantum information theory," *Probl. Inform. Transm.*, vol.36, pp.25–34, 2000. (Originally appeared in LANL archive math-ph/0003002.)

[2] S. Arimoto, "An algorithm for calculating the capacity of an arbitrary discrete memoryless channel, "*IEEE Trans. Inform. Theory*, vol.18, pp.14–20, 1972.

[3] C.H. Bennett, C.A. Fuchs, and J.A. Smolin, "Entanglement-enhanced classical communication on a noisy quantum channel," *Quantum Communication, Computing and Measurement,* eds. O. Hirota, A.S. Holevo, and C.M. Caves, pp.79–88, Plenum, 1997. (Originally appeared in LANL archive quant-ph/9611006.)

[4] R. Blahut, "Computation of channel capacity and rate distortion functions," *IEEE Trans. Inform. Theory*, vol.18, pp.460–473, 1972.

[5] D. Bruss, L. Faoro, C. Macchiavello, and G.M. Palma, "Quantum entanglement and classical communication through a depolarising channel," *J. Mod. Opt.*, vol.47, pp.325–332, 2000. (Originally appeared in LANL archive quant-ph/9903033.)

[6] T.M. Cover and J.A. Thomas, "*Elements of Information Theory*," Wiley, 1991.

[7] A. Fujiwara and H. Nagaoka, "Operational capacity and pseudoclassicality of a quantum channel," *IEEE Trans. Inform. Theory*, vol.44, pp.1071–1086, 1998.

[8] A. Fujiwara and P. Algoet, "One-to-one parametrization of quantum channels," *Phys. Rev. A*, vol.59, pp.3290–3294, 1999.

[9] A.S. Holevo, "Bounds for the quantity of information transmitted by quantum communication channel," *Probl. Inform. Transm.*, vol.9, no.3, pp.177–183, 1973.

[10] A.S. Holevo, "On the capacity of quantum communication channel," *Probl. Inform. Transm.*, vol.15, no.4, pp.247–253, 1979.

[11] A.S. Holevo, "The capacity of the quantum channel with general signal states," *IEEE Trans. Inform. Theory*, vol.44, pp.269–273, 1998. (Originally appeared in LANL archive quant-ph/9611023.)

[12] A.S. Holevo, "Coding theorems for quantum channels," LANL archive quant-ph/9809023.

[13] C. King and M.B. Ruskai, "Minimal entropy of states emerging from noisy quantum channels," *IEEE Trans. Inform. Theory*, vol.47, pp.192–209, 2001. (Originally appeared in LANL archive quant-ph/9911079.)

[14] C. King and M.B. Ruskai, "Capacity of quantum channels using product measurements," *J. Math. Phys.*, vol.42, pp.87–98, 2001. (Originally appeared in LANL archive quant-ph/0004062.

[15] C. King, "Maximization of capacity and l_p norms for some product channels," LANL archive quant-ph/0103086.

[16] C. King, "Additivity for a class of unital qubit channels," LANL archive quant-ph/0103056.

[17] K. Kraus, "*States, Effects, and Operations: Fundamental Notions of quantum Theory*," Springer, 1983.

[18] H. Nagaoka, "Algorithms of Arimoto-Blahut type for computing quantum channel capacity," *Proc. of 1998 IEEE International Symposium on Information Theory*, p.354, 1998.

[19] H. Nagaoka and S. Osawa, "Theoretical basis and applications of the quantum Arimoto-Blahut algorithms," *Proc. of the second QIT*, pp.107–112, 1999.

[20] S. Osawa and H. Nagaoka, "Numerical experiments on the quantum channel capacity when input states can be entangled," *Proc. of the 22nd Symposium on Information Theory and Its Applications*, pp.387–390, 1999.

[21] B. Schumacher and M.D. Westmoreland, "Sending classical information via noisy quantum channels," *Phys. Rev. A*, vol.56, pp.131–138, 1997.

[22] B. Schumacher and M.D. Westmoreland, "Relative entropy in quantum information theory," LANL archive quant-ph/0004045.

[23] W.F. Stinespring, "Positive functions on C^*-algebras," *Proc. Amer. Math. Soc.*, vol.6, pp.211–216, 1955.

[24] V. Vedral and M.B. Plenio, "Entanglement measures and purification procedures," *Phys. Rev. A*, vol.57, pp.1619–1633, 1998.